Intrusion Discovery using SVM: A Survey

Naseer Ahmed Shah[1], Jyoti Arora[2],
[1]*M.Tech Student, Desh Bhagat University,Mandi Gobindgarh*
[2] *Assistant Professor, Desh Bhagat University,Mandi Gobindgarh*
(E-mail: pnz101786@gmail.com)

Abstract— Intrusion is generally defined as a successful attack on a network or system. In a technical report on the practice of Intrusion Recovery. The intruder carries out an attack with a specific objective in mind. From the perspective of an administrator responsible for maintaining a system, an attack is a set of one or more events that may have one or more security consequences. An Intrusion Detection is an important in assuring security of network and its different resources. Intrusion detection attempts to detect computer attacks by examining various data records observed in processes on the network. Recently data mining methods have gained importance in addressing network security issues, including network intrusion detection. Intrusion detection systems aim to identify attacks with a high detection rate and a low false positive. Here, we are going to study Intrusion Detection System using data mining technique: Support Vector Machine (SVM). Support vector machine-based intrusion detection methods are increasingly being researched because it can detect novel attacks.

Keywords—*Intrusion Detection; SVM; Classification, Network Security.*

I. INTRODUCTION

The Vast application of computer networks, the quantity of attacks, hacking tools and intrusive methods have developed widely. One of the way of dealing with suspicious actions is by utilizing an intrusion detection system (IDS). By investigation of many records on the network [1] [2].

The security of a computer system is vulnerable when an intrusion takes place. An intrusion can be defined as any action done that harms the integrity, confidentiality or availability of the system. There are some intrusion prevention techniques which can be used to prevent computer systems as a first line of defense. A firewall is also one of it. But only intrusion prevention is not enough. As systems become more complex, there are always exploitable weaknesses in the systems due to design and programming errors, or various penetration techniques. Therefore Intrusion detection is required as another measure to protect our computer systems from such type of vulnerabilities [3].

Intrusion detection approaches are commonly divided into two categories: signature or misuse detection and anomaly detection [7] [8]. Misuse detection is the ability to identify intrusions based on a known pattern for the malicious activity .These known patterns are referred to as signatures. The idea of misuse detection is to establish a pattern or a signature form so that the same attack can be detected. Thus, the main drawback of misuse detection is it cannot detect new types of attacks. The IDS has a pattern database that includes signatures of possible attacks. If the system matches the data with the attack pattern, the IDS regards it as an attack.

Consequently, misuse detection provides a low false positive rate (the rate of misclassified normal behavior).Anomaly detection is the attempt to identify malicious traffic based on deviations from established normal network traffic patterns. Anomaly detection requires storage of normal usage behavior and operates upon audit data generated by the operating system. With the anomaly detection approach, one represents patterns of normal behavior, with the assumption that an intrusion can be identified based on some deviation from this normal behavior. When such a deviation is observed, an intrusion alarm is produced.

Anomaly detection is capable of catching new attacks, yet it suffers a higher false positive rate. Intrusion detection is categorized into two, host based and network based approaches. Network intrusion detection system (NIDS) [10] detects intrusions by continuously monitoring network traffic by connecting to network hub or switch which is configured for port mirroring, or network tap.

NIDS uses sensors to capture all network traffic and to monitor individual packets to identify whether it is normal or attack. Network based IDS is installed on network elements like routers to monitor the network traffic. In network-based (NIDS), the packets are collected from the network. An example of a NIDS is Snort.

A host-based intrusion detection system (HIDS) [9] is an intrusion detection system that monitors and analyzes the internals of a computing system as well as the network packets on its network interfaces. Host based IDS is installed on each system for monitoring of malicious activities locally .Host-based intrusion detection system (HIDS) uses agent as a sensor on a host that identifies intrusions by analyzing system calls, application logs, file-system modifications (binaries, password files, etc.) and other host activities and state. OSSEC is an example for Host based intrusion detection system.

Intrusion prevention techniques, such as user authentication and information protection via encryption have been used to protect computer systems as a first line of defense. Intrusion prevention alone is not sufficient because as systems become more complex, there are always exploitable weaknesses in the systems due to design and programming errors, or various "socially engineered" penetration techniques. Intrusion detection is therefore needed as another wall to protect computer systems. IDS system is only detect the intrusion with the help of different classification algorithm. The main functionality in intrusion system is performed by classification

algorithm. There are several algorithm used with IDS such as ANN, FFNN, Fuzzy Logics, SVM and genetic algorithm with SVM. Here we are studying the performance and use of SVM for Intrusion Detection.

II. LITERATURE SURVEY

There are many attempts have been done for SVM to get maximum accuracy in IDS. Some of these techniques are studied below.

In 2011, Horng, Shi-Jinn, et al. proposed an SVM based intrusion detection system, which used hierarchical clustering algorithm, leave one out, and the SVM technique. The hierarchical clustering algorithm provided the SVM with fewer, abstracted, and higher-qualified training instances that are derived from the KDD Cup 1999 training set. It was able to greatly minimize the training time, and improve the performance of SVM. The simple feature selection procedure (leave one out) was applied to eliminate unimportant features from the training set so the obtained SVM model could classify the network traffic data more accurately [1].

In 2012, Gaspar, Paulo, Jaime Carbonell, and José Luís Oliveira et al. gave the review on strategies that are used to improve the classification performance in term of accuracy of SVMs and perform some experimentation to study the influence of features and hyper-parameters in the optimization process, using kernels function. Huang et al provide a study on the joint optimization of C and g parameters (using the RBF kernel), and feature selection using Grid search and genetic algorithms [2].

In 2014, Ahmad, Iftikhar, et al. proposed a genetic algorithm to search the genetic principal components that offers a subset of features with optimal sensitivity and the highest discriminatory power. The support vector machine (SVM) is used for classification. The results show that proposed method enhances SVM performance in intrusion detection [3].

In 2008, Zhou, Jianguo, et al. Proposed system a Culture Particle Swarm Optimization algorithm (CPSO) used to optimize the parameters of SVM. By using the colony aptitude of particle swarm and the ability of conserving the evolving knowledge of the culture algorithm, this CPSO algorithm constructed the population space based on particle swarm and the knowledge space. The proposed CPSO-SVM model that can choose optimal values of SVM parameters was test on the prediction of financial distress of listed companies in China [5].

In 2011, Kolias, Constantinos, Georgios Kambourakis, and M. Maragoudakis et al. suggested that the RBF has certain parameter that affects the accuracy. PSO is used along with RBF artificial neural network it will improve the accuracy. If it is used in IDS it will improves the accuracy of classification [6].

Furthermore, Heba F. Eid, Ashraf Darwish, Aboul Ella Hassanien, and AjithAbrahamwe effectively introduced intrusion detection system by using Principal Component Analysis (PCA) with Support Vector Machines (SVMs) as an approach to select the optimum feature subset [25] they verified the effectiveness and the feasibility of the proposed IDS system by several experiments on NSL-KDD dataset.

J.F Joseph, A. Das, B.C. Seet in their paper proposed an autonomous host-based ID for detecting sinking behavior in an ad hoc network [26]. The proposed detection system uses a cross-layer approach to maximize detection accuracy. To further maximize the detection accuracy SVM is used for training the detection model.

However, SVM is computationally expensive for resource-limited ad hoc network nodes. Hence, the proposed IDS preprocess the training data for reducing the computational overhead incurred by SVM. Number of features in the training data is reduced using predefined association functions. Also, the proposed IDS uses a linear classification algorithm, namely Fischer Discriminants Analysis (FDA) to remove data with low-information content (entropy). The above data reduction measures have made SVM feasible in ad hoc network nodes.

T. Shon, Y. Kim, C. Lee and J. Moon in their paper proposed a Machine Learning Model using a modified Support Vector Machine (SVM) that combines the benefits of supervised and unsupervised learning [27]. Moreover, a preliminary feature selection process using GA is provided to select more appropriate packet fields.

Peddabachigari, A. Abraham, C. Grosan conducted an empirical investigation of SVM and Decision Tree, in which they analyzed their performance as standalone detectors and as hybrids [28]. Two hybrids models were examined, a hierarchical model (DT-SVM), with the DT as the first layer to produce node information for the SVM in the second layer, and an ensemble model comprising the standalone techniques and the hierarchal hybrid. For the ensemble approach, each technique is given a weight according to detection rate of each particular attack type during training. Thereafter, when the system is tested, only the technique with the largest weight for the respective attack prediction is chosen to output the classification. The approaches were tested on the KDD Cup ’99 data set.

R. C. Chen, K.F Cheng and C. F Hsieh in their paper used RST (Rough Set Theory) and SVM (Support Vector Machine) to detect intrusions [29]. First, RST is used to preprocess the data and reduce the dimensions. Next, the features selected by RST are sent to SVM model to learn and test respectively. The method is effectively decreased the space density of data.

Kyaw The tKhaingin his paper proposed an enhanced SVM Model with a Recursive Feature Elimination (RFE) and kNearest Neighbor (KNN) method to perform a feature ranking and selection task of the new model [30].

III. SUPPORT VECTOR MACHINE

In this paper the different approaches for the SVM are discussed for intrusion detection. Based on the structural risk minimization principle, SVM is a new machine learning method presented by Vapnik (1995) [13].

Generalization ability of SVM is obviously superior to other traditional learning methods. This basic SVM deals with two-class problems, known as Binary classification problems in which the data are separated by a hyper plane defined by a number of support vectors. Support vectors are a subset of training data used to define the boundary between the two classes. Each instance in the training set contains one "target value" (class labels: Normal or Attack) and 41 features. The goal of SVM is to produce a model which predicts target value of data instance in the testing set which consists of only features. To achieve this goal, we have used Radial Basis Function (RBF) kernel functions [29, 31] available with SVM. In situations where SVM cannot separate two classes, it solves this problem by mapping input data into high-dimensional feature spaces using a kernel function [14, 33].

In high dimensional space it is possible to create a hyper plane that allows linear separation (which corresponds to a curved surface in the lower- dimensional input space). Accordingly, the kernel function plays an important role in SVM. The kernel functions can be used at the time of training of the classifiers which selects support vectors along the surface of this function. SVM classify data by using these support vectors that outline the hyper plane in the feature space. In practice, various kernel functions can be used, such as linear, polynomial or Gaussian. The SVM is already known as the best learning algorithm for binary classification [11] [15] [16].

However, it is not the reason that we have chosen SVM. The most significant reason we chose the SVM is because it can be used for either supervised or unsupervised learning. The SVM, originally a type of pattern classifier based on a statistical learning technique for classification and regression with a variety of kernel functions [7, 19], has been successfully applied to a number of pattern recognition applications [15]. Recently, it has also been applied to inform security for intrusion detection [17] [8].

Another positive aspect of SVM is that it is useful for finding a global minimum of the actual risk using structural risk minimization, since it can generalize well with kernel tricks even in high-dimensional spaces under little training sample conditions. The SVM can select appropriate setup parameters because it does not depend on traditional empirical risk such as neural networks. In the case of supervised SVM learning, it has relatively fast processing and high detection performance when compared to existing artificial neural networks and the unsupervised SVM , as shown in[24][25].

However, one of the main disadvantages of the supervised method is that it requires labelled information for efficient learning. Moreover, it cannot deal with the KDD99 Dataset Pre Processing SVM Train SVM Test Result Analysis relationship between consecutive variations of learning inputs without additional pre-processing. Therefore, Taeshik Shon and Jongsub Moon have proposed the real time intrusion detection system using Enhanced SVM, which combines soft margin SVM using supervised learning and one-class SVM approach using the unsupervised learning. The enhanced SVM approach inherits the advantages of both SVM approaches, namely high performance and unlabeled capability. The SVM is generally used as a supervised learning method.

Vapnik proposed the initial idea of SVM for the separable case (hard margin SVM) in which the positive and negative samples can be definitely separated by a unique optimal hyper plane with the largest margin. However, this algorithm will find no feasible solution when applied to the non-separable case. Cortes and Vapnik extended this idea to the no separable case (soft margin SVM or the so called standard SVM) by introducing positive slack variables I=1, l .In order to decrease misclassified data, a supervised SVM approach with a slack variable is called soft margin SVM. Additionally, single class learning for classifying outliers can be used as an unsupervised SVM. After considering both SVM learning schemes, an Enhanced SVM approach is proposed.

IV. CONCLUSION

The Network Intrusion Detection System is the latest technology which is an important technology of the Network Security Nowadays there are many advanced technology for the Intrusion Detection Systems. The major reasons for using SVMs in intrusion detection system is speed: as real-time performance is of primary importance to IDSs, any classifier that can potentially run "fast" is worth considering and the another reason is scalability, SVMs are relatively insensitive to the number of data points and the classification complexity does not depend on the dimensionality of the feature space, so they can potentially learn a larger set of patterns and thus be able to scale better than neural networks. Also SVM provides a standard mechanism to fit the surface of the hyper plane to the data by utilizing the kernel function.

Among the variety of Intrusion detection approaches, the Support Vector Machine (SVM) is known to be one of the best machine learning algorithms to classify abnormal behavior. This research has the following two contributions. First, this paper provides a review on current trends in intrusion detection using SVM together with a study on technologies implemented by some researchers in this research area. Second it proposes a novel approach SVM to select best feature for detecting intrusion.

REFERENCES

[1] Anderson, J.P," Computer Security Threat Monitoring and Surveillance," Technical Report, Vol.3, pp.234- 267, 1980.

[2] Yang Li, Li Guo, "An Active Learning Based TCM-KNN Algorithm for Supervised Network Intrusion Detection," Computers & Security, vol.26, pp.459-467, 2007.

[3] Ahmad, Iftikhar, et al. "Enhancing SVM performance in intrusion detection using optimal feature subset selection based on genetic principal components." Neural Computing and Applications 24.7-8 (2014): 1671-1682.

[4] Hashem, Soukaena Hassan. "Efficiency of Svm and Pca to Enhance Intrusion Detection System." Journal of Asian Scientific Research 3.4 (2013): 381-395.

[5] Zhou, Jianguo, et al. "The study of SVM optimized by Culture Particle Swarm Optimization on predicting financial distress." Automation and Logistics, 2008. ICAL 2008. IEEE International Conference on. IEEE, 2008.

[6] Kolias, Constantinos, Georgios Kambourakis, and M. Maragoudakis. "Swarm intelligence in intrusion detection: A survey." computers & security 30.8 (2011): 625-642.

[7] Lee W and Stolfo S., "Data Mining techniques for intrusion detection", In: Proc. of the 7th USENIX security symposium, San Antonio, TX, 1998.

[8] Dokas P, Ertoz L, Kumar V, Lazarevie A, Srivastava J, and Tan P., "Data Mining for intrusion detection", In: Proc. of NSF workshop on next generation data mining, 2002.

[9] De Boer P., Pels M. "Host-Based Intrusion Detection Systems".Availablehttp://staff.science.uva.nl/~delaat/snb -2004-2005/p19/report.pdf.

[10] Scarfone K., Mell P. "Guide to Intrusion Detection and Prevention Systems". Available at http://csrc.nist.gov/publications/nistpubs/80094/SP80094 .pdf, 2007.

[11] C. Cortes and V. Vapnik, "Support-vector network," Machine Learning, vol. 20, pp. 273–297, 1995

[12] [8] S. Mukkamala, G.1. Janoski, A.H. Sung. Intrusion Detection Using Neural Networks and Support Vector Machines. In Proceedings of IEEE International Joint Conference on Neural Networks, Vol 2, Honolulu, 2002.5, pp. 1702-1707.

[13] V. N. Vapnik. The nature of statistical learning theory. Springer Verlag, New York. NY, 1995

[14] C.J.C. Burges, A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, vol 2(2), Springer US, 1998, pp.121-167.

[15] K.-P. Lin and M.-S. Chen, "Efficient kernel approximation for large-scale support vector machine classification," in Proceedings of the Eleventh SIAM International Conference on Data Mining, 2011, pp. 211– 222

[16] H. Byun, S.W. Lee, A survey on pattern recognition applications of support vector machines, International Journal of Pattern Recognition and Artificial Intelligence 17 (2003) 459–486

[17] Amit Konar, Uday K. Chakraborty, Paul P. Wang, Supervised learning on a fuzzy Petri net, Information Sciences 172 (2005) 397–416

[18] B. Scho ¨lkopf, estimating the support of a high dimensional distribution, Neural Computation 13 (2001) 1443–147.

[19] T. Joachims, Estimating the Generalization Performance of an SVM efficiently, in: Proc. the Seventeenth International Conference on Machine Learning, San Francisco, CA, 2000, pp. 431–438

[20] B.V. Nguyen, An Application of Support Vector Machines to Anomaly Detection, CS681 (Research in Computer Science – Support Vector Machine) report, 2002

[21] S. Dumais, H. Chen, Hierarchical classification of Web content, in: Proc. The 23rd annual international ACM SIGIR conference on Research and development in information retrieval, Athens, Greece, 2000, pp. 256–263

[22] S. S. Keerthi and C.-J. Lin, "Asymptotic behaviors of support vector machines with Gaussian kernel," Neural Computation, vol. 15, no. 7, pp. 1667–1689, 2003.

[23] K. Crammer and Y. Singer. On the algorithmic implementation of multiclass kernel-based vector machines. Journal of Machine Learning Research, 2:265–292, 2001.

[24] N. Cristiani and J. Shawe-Taylor. An Introduction to Support Vector Machines and other kernel-based learning methods. Cambridge, Cambridge University Press, 2000.

[25] Heba F. Eid, Ashraf Darwish, Aboul Ella Hassanien, and Ajith Abraham, (2010) Principle Components Analysis and Support Vector Machine based Intrusion Detection System, IEEE.

[26] J.F Joseph, A. Das,B.C. Seet, (2011) Cross-Layer Detection of Sinking Behavior in Wireless Ad Hoc Networks Using SVM and FDA. IEEE Transaction on dependable and secure computing, Vol. 8, No. 2, MarhApril 2011.

[27] T.Shon, Y. Kim, C.Lee and J.Moon,(2005), A Machine Learning Framework for Network Anomaly Detection using SVM and GA, Proceedings of the 2005 IEEE.

[28] Sandya Peddabachigari, Ajith Abraham, Crina Grosan, Johansson Thomas (2005). Modeling Intrusion Detection Systems using Hybrid Intelligent Systems. Journal of Network and Computer Applications.

[29] R.C. Chen, K.F Cheng and C. F Hsieh (2009), using support vector machine and rough set for network intrusion system.

[30] KyawThetKhaing (2010), Recursive Feature Elimination (RFE) and k-Nearest Neighbor (KNN) in SVM.